Emotion
2006, Vol. 6, No. 4, 663–669

1528-3542/06/$12.00 DOI: 10.1037/1528-3542.6.4.663

BRIEF REPORTS

Exploring the Validity of the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) with Established Emotions Measures

Richard D. Roberts and Ralf Schulze
Center for New Constructs, Educational Testing Service

Kristin O'Brien and Carolyn MacCann
University of Sydney

John Reid
Macquarie University

Andy Maul
University of California at Berkeley

Emotions measures represent an important means of obtaining construct validity evidence for emotional intelligence (EI) tests because they have the same theoretical underpinnings. Additionally, the extent to which both emotions and EI measures relate to intelligence is poorly understood. The current study was designed to address these issues. Participants ($N = 138$) completed the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT), two emotions measures, as well as four intelligence tests. Results provide mixed support for the model hypothesized to underlie the MSCEIT, with emotions research and EI measures failing to load on the same factor. The emotions measures loaded on the same factor as intelligence measures. The validity of certain EI components (in particular, Emotion Perception), as currently assessed, appears equivocal.

Keywords: emotional intelligence, emotion perception, validity, factor analysis

Emotional intelligence (EI) has been defined as the ability to perceive emotions, to use emotions to facilitate performance on cognitive tasks, to understand emotions, and to regulate or manage emotions effectively (Mayer & Salovey, 1997). It has been argued that if the meanings of the words *emotion* and *intelligence* are to be preserved, the term *emotional intelligence* should combine them in an effective manner (Matthews, Zeidner, & Roberts, 2002). In particular, it should combine the ideas that emotions can make thinking more intelligent and that one can think intelligently about emotions. However, to date, there has been a paucity of research linking EI to other emotions constructs and relatively little research linking EI to cognitive ability constructs (e.g., fluid and crystallized intelligence). The primary aim of the present investigation was to provide new evidence on these links and, in the process, establish validity evidence for an often-used measure of EI.

Elsewhere, we have drawn attention to a series of problems concerning the measurement of EI, particularly those based on self-report techniques (e.g., Matthews, Roberts, & Zeidner, 2004). The more compelling approach, involving performance-based assessment, has been conducted almost exclusively using the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT; Mayer, Salovey, & Caruso, 2002). Items of the MSCEIT are grouped into eight tasks (e.g., Faces, Pictures), with four branches (Emotions Perception, Facilitating Cognition through Emotions, Understanding Emotions, and Managing Emotions), which combine to form two areas (Experiential and Strategic). A general EI factor is also hypothesized at the apex of this hierarchy (Mayer et al., 2002).

An issue of utmost importance in the field of EI is the validity of the MSCEIT and whether it can meet comparable criteria to that found for other intelligence constructs (Austin & Saklofske, 2005; Matthews et al., 2002; Roberts, Schulze, Zeidner, & Matthews, 2005; Wilhelm, 2005). Elsewhere, this debate has centered on whether EI meets the following criteria: (a) operationalizability; (b) objective scoring; (c) relations with established intelligence constructs; (d) unique variance; and (e) age trends (Mayer, Salovey, Caruso, & Sitarenios, 2003; Roberts, Zeidner, & Matthews, 2001). While studies have addressed these specific criteria (cf. Mayer et al., 2003; MacCann, Roberts, Matthews, & Zeidner, 2004; Roberts et al., 2001), others have focused on relations with additional measures of EI, personality factors, and performance criteria (Brackett & Mayer, 2003; Day & Carroll, 2004; Kafetsios, 2004; Lopes, Salovey, & Straus, 2003). A notable omission from

Richard D. Roberts and Ralf Schulze, Center for New Constructs, Educational Testing Service; Kristin O'Brien and Carolyn MacCann, Department of Psychology, University of Sydney; John Reid, Department of Psychology, Macquarie University; Andy Maul, Graduate School of Education, University of California at Berkeley.

The views expressed here are those of the authors and do not reflect on the Educational Testing Service. This research was conducted as part of the third and fourth authors Honors dissertation under the supervision of the first author while at the School of Psychology, The University of Sydney. We are grateful to Paul Cruz, Walter Emmerich, Matthew Ventura, the reviewers of the manuscript, and colleagues at the Individual Differences Computerized Assessment Unit at The University of Sydney for comments on earlier drafts of this paper.

Correspondence concerning this article should be addressed to Richard D. Roberts, ETS (Center for New Constructs), Rosedale Road MS 16-R, Princeton, NJ 08541. E-mail: RRoberts@ets.org

the list of constructs that would seem pivotal in obtaining construct validity evidence appears those tied to emotions processes. Thus, in addition to exploring criteria (c) and (d) listed above for an intelligence, the present study sought to investigate relations with established emotions constructs and corresponding measures.

Is the MSCEIT Measuring Emotional Processes and/or Components?

The study of emotions and how people interpret them is not new. Experimental as well as correlational approaches to the measurement and study of emotions have a long tradition outside of the ostensibly younger field of EI research (see Matthews et al., 2002). Hence, it seems reasonable that there are measurement approaches in emotions research that might profitably be used to validate performance-based EI assessments. In the present paper, we focus on assessment procedures to measure perception of emotions in faces and voices from emotions research as a means of providing construct validity evidence for the MSCEIT and, in particular, subtests of this instrument that measure Emotion Perception.

Examples of cognitive-affective research tools for examining the ability to perceive emotions in facial expressions include Ekman's standardized database of facial expressions (Ekman, 1973) and the Japanese and Caucasian Brief Affect Recognition Test (JACBART; Matsumoto et al., 2000), among others. Instruments have also been developed to measure emotion recognition ability pertaining to tone-of-voice: the Index of Vocal Emotion Recognition (Vocal-I; Scherer, Banse, & Wallbott, 2001; Scherer, in press) and the Diagnostic Assessment of Non Verbal Affect—Adult Paralanguage (DANVA2-AP; Baum & Nowicki, 1998). At a conceptual level, the ability to recognize emotions is essentially identical to the first branch assessed by the MSCEIT: Emotion Perception. Hence, it is hypothesized that performance on tasks from emotions research, which assess the ability to recognize or decode emotions, are indicators of the same latent variable as the first branch indicators of the MSCEIT (i.e., Faces and Pictures).

It is noteworthy that, in contrast to the tasks used in EI research, there can usually be found a considerable body of empirical evidence for the reliability and validity of the measures originating from emotions research (Matsumoto et al., 2000; O'Sullivan, 1982). Conceivably, this feature allows them to provide construct validity evidence for the relatively newer EI tests that have recently been developed. In terms of validating the Emotion Perception component of the MSCEIT, we might thus hypothesize:

> Tests from emotions research and Emotion Perception from the MSCEIT are indicators of the same latent variable.

Is the MSCEIT Measuring an Intelligence?

The available evidence suggests that EI relates to established intelligence constructs. For example, Lopes et al. (2003) found a significant relationship between WAIS Verbal IQ and the Understanding Emotions branch score ($r = .39$). Other studies have found relationships with SAT scores, college grades, and college admissions indices (e.g., Brackett & Mayer, 2003; see also Davis & Kraus, 1997). In general, it appears that EI relates to acculturated (or crystallized) intelligence, and those relations are strongest for the third branch (i.e., Understanding; see Roberts et al., 2005).

In the present study, we wished to test this proposition further. Under the framework of Gf-Gc theory, crystallized intelligence (Gc) is a measure of acquired knowledge and skill, while fluid intelligence (Gf) is a measure of the ability to solve novel problems (Carroll, 1993). Markers of Gf and Gc were included in the present investigation in an attempt to relate EI and measures from emotions research to fluid, as well as crystallized, intelligence. A review of the pertinent literature suggests:

> Components of EI from the MSCEIT and Gc are moderately positively correlated (strong for Understanding). Correlations of EI components and Gf are positive and weak to moderate.

In addition to the preceding hypotheses, we also explore the relationships between selected tasks from emotions research with intelligence measures.

Method

Participants

Undergraduate psychology students ($N = 138$) from the University of Sydney participated to fulfill course requirements. Among 136 participants, for whom demographic data was available, 87 were female and the mean age was 20.7 ($SD = 6.32$) years. In all other respects, the demographics of the sample mirrored closely that of the Australian population.

Test Descriptions

Participants completed a biographical questionnaire, the MSCEIT, emotions measures (JACBART, Vocal-I),[1] and cognitive ability tests (Vocabulary, Esoteric Analogies, Swaps, and Matrices). Each task was presented as part of a computer-based battery.

Emotional Intelligence

The MSCEIT consists of two tests for each of the four branches: (a) Perceiving Emotions (Faces, Pictures); (b) Facilitating Thoughts (Facilitation, Sensations); (c) Understanding Emotions (Blends, Changes); and (d) Managing Emotions (Management, Relations). A detailed description of these tests may be found in Mayer et al. (2002; see also Matthews et al., 2002). Tests were scored according to the consensus method. The procedure consists of awarding an option the score equal to the proportion of a prior screening sample who selected that option (e.g., if 58% of the screening sample chooses "A", then the score awarded for that option is .58; Legree, Psotka, Tremble, & Bourne, 2005). In this case, we used the

[1] In addition to these two tasks from emotions research, the Emotional Stroop (ES) task (McKenna & Sharma, 1995) was also included in the set of measures. This measurement paradigm is widely used in clinical psychology and seems to be particularly interesting for EI research (Coffey, Berenbaum, & Kerns, 2003). It can be argued that the ES reflects individual differences in the ability to intentionally protect thinking processes from being impaired by emotions, which seems to be closely and negatively related to the EI component Facilitating Cognition through Emotions. Therefore, we hypothesized that Emotional Stroop and Using Emotions to Facilitate Cognition tests from the MSCEIT are indicators of the same latent variable. Since the expected interference effect in the ES task (see McKenna & Sharma, 1995) did not occur, the hypothesis could not be tested and the results for this task are not included in this paper.

current sample of Australian students to determine weights rather than the standard weights of the American sample.[2]

Emotions Measures

Index of Vocal Emotion Recognition (Vocal-I). In this task, participants were presented with two meaningless sentences ("Hat sundig pron you venzy" and "Fee gott laish jonkill gosterr"), which were recited by male and female radio actors to portray joy, sadness, fear, anger, or no emotion. For each of the 30 items, participants were required to choose the emotion present (Scherer et al., 2001).

Japanese and Caucasian Brief Affect Recognition Test (JACBART). This test consisted of 56 Japanese or Caucasian faces portraying one of seven emotions: happiness, contempt, disgust, sadness, anger, surprise, and fear. Each stimulus, sandwiched between a backward and forward mask that shows a neutral face, is presented for 200 msec. The participants' task was to determine the emotion portrayed in each respective face (Matsumoto et al., 2000).

Cognitive Ability Measures

Two measures (Vocabulary and Esoteric Analogies) included to assess Gc were taken from the Gf/Gc Quickie Test Battery, while Swaps and a Matrices test were expected to be markers of Gf (see Roberts & Stankov, 1999). Capsule descriptions of these tests follow.

Vocabulary. Participants answered 30 questions related to the meaning of words. For example, "What is the meaning of excerpt?" (Responses: "Accept, *Extract*, Curtail, Deprive").

Esoteric Analogies. Participants completed 23 analogies having the following form: "Statue is to shape as song is to?" (Responses: "Beauty, Piano, *Tune*, Note").

Swaps. Participants were required to mentally swap the order of letters in a three-letter string. All swaps had to be applied to the result of previous swaps. For example, J L K; (a) swap 1 and 3; (b) swap 1 and 2 (Answer = L K J). The number of swaps (a measure of difficulty) ranged from 1 to 4. Each participant completed 4 items at each difficulty level, such that there were 16 items in total.

Matrices. Participants were required to identify consistencies in a pattern of designs going down and across a 3×3 matrix. The task was to select a design, from a set of 6–8 options that fit into an empty cell in the 3×3 matrix according to the perceived rule in the matrix (12 items).

Procedure

Groups of up to 14 participants took the test battery in a laboratory environment. Approximately 1.5 hours were needed to complete the battery. Sydney University's Human Research Ethics committee approved the study.

Results

Descriptive Statistics

The means, standard deviations, and estimates of reliability (Cronbach's α) for the MSCEIT, intelligence, and emotions measures are presented in Table 1. The means from the MSCEIT obtained in the present sample are close to those obtained in the original study conducted by Mayer et al. (2002) based on an American sample. The only difference larger than half a standard deviation (see *d* values) is observed for Management subtest. Means for the tests from emotions research were also similar to previous studies conducted by the test authors; and statistics for the cognitive ability tests were comparable to our previous studies, although slightly higher in the present study for Gc measures. Correlations between all instruments are presented in the Appendix.

Exploratory and Confirmatory Factor Analyses

An exploratory maximum likelihood factor analysis (EFA) of the MSCEIT, Gc, Gf, and other measures was conducted with CEFA (Browne, Cudeck, Tateneni, & Mels, 2004) to explore the factorial structure of these tests. The results of a parallel analysis (O'Connor, 2000) suggested the extraction of three factors. These three factors were extracted and then rotated using an oblique procedure (CF-Varimax; see Browne, 2001).

The model fit very well ($\chi^2 = 43.19$, $df = 52$, $p = .80$). Factor loadings, presented in Table 2, suggest that Factor 1 corresponds to Experiential EI, while Factor 2 corresponds to Strategic EI. Loadings on Factor 3 are significant for Blends, Changes (Emotional Understanding), Vocabulary, Analogies, Swaps, Vocal-I, and JACBART. Hence, we interpret this factor as an intelligence factor, which subsumes Vocal-I and JACBART, suggesting that these measures of emotional processes at least partially reflect intelligence (see also Scherer, in press). Both Gf and Gc indicators load on this intelligence factor, which precludes a differentiated assessment of the relationships between EI factors and Gf or Gc, respectively. Note that the two subtests Changes and Blends from the MSCEIT have cross-loadings on Factor 3 and are therefore complex. This is consistent with previous findings showing that measures of the Understanding branch are related to intelligence factors (Roberts et al., 2005). All correlations between the three factors were positive, small to medium in size, and significant. This conforms to the expectations stated above, though the pattern of relationships does not reflect the expected strong association between Experiential EI and emotions research measures. The Matrices test, which was expected to load on a Gf factor, did not load significantly on any of the factors and was dropped from subsequent analyses.

A confirmatory factor analysis (CFA) was run to provide additional information on the three-factor model by constraining nonsignificant loadings in the EFA to zero. Table 2 shows the factor loadings, factor correlations, and fit measures. Overall, the results largely overlap with those from the previous EFA. All factors correlate positively with each other, where the correlation between the Experiential EI factor and the third factor is lowest among these correlations.

With respect to the hypotheses, the most important results are (a) tests from emotions research and Emotion Perception from the MSCEIT are *not* indicators of the same latent variable, and (b) strategic components of EI from the MSCEIT and a factor of intelligence, which combined Gf and Gc indicators, were moderately positively correlated.

[2] This scoring system is justifiable on the following grounds: (a) the correlation between consensus-weights derived from disparate samples tends to exceed .90 (Zeidner, Shani-Zinovich, Matthews, & Roberts, 2005); (b) the researcher is free to determine reliability using this approach (using American sample weights provided by the test publisher, this is not generally the case); and (c) possible cultural confounds are likely minimized using Australian weights.

Table 1
Comparison of Means, Standard Deviations, and Reliabilities for Tests Used in the Present Study (N = 138) With Comparative Data

Measures	*M*		*SD*			α	
	Current	CD	Current	CD	*d*	Current	CD
MSCEIT							
Branch 1: Emotion Perception							
Faces	.48	.51	.07	.12	−0.26	.68	.81
Pictures	.48	.53	.08	.13	−0.39	.80	.88
Branch 2: Emotion Facilitation							
Facilitation	.48	.52	.08	.11	−0.37	.61	.64
Sensations	.40	.44	.06	.09	−0.45	.38	.65
Branch 3: Emotional Understanding							
Changes	.61	.57	.08	.10	0.40	.52	.70
Blends	.55	.53	.08	.10	0.20	.50	.66
Branch 4: Emotion Management							
Management	.37	.44	.05	.09	−0.79	.52	.69
Relations	.42	.46	.08	.11	−0.37	.51	.67
Cognitive Ability Measures							
Vocabulary	.77	.73	.14	.14	0.29	.71	.60
Esoteric Analogies	.66	.62	.16	.14	0.27	.70	.78
Swaps	.71	.70	.22	.20	0.05	.73	.66
Matrices	.46	.43	.18	.20	0.16	.70	.73
Emotions Measures							
Vocal-I	.66	.67	.10	.10	−0.10	.45	.43
JACBART	.72	.68	.11	.13	0.34	.73	.82

Note. Data for the MSCEIT was consensus scored; cognitive ability and emotion measures were veridically scored. MSCEIT = Mayer-Salovey-Caruso Emotional Intelligence Test; JACBART = Japanese and Caucasian Brief Affect Recognition Test; Current = Data from the present study; CD = Comparative data from Mayer et al. (2002) for MSCEIT ($N = 2112$); Markham (2002) for cognitive ability measures ($N = 146$); Matsumoto et al. (2000) for the JACBART ($N = 89$, Study 2); and MacCann (2005) for the Vocal-I ($N = 121$), d = standardized mean difference between the current study and comparative data, where the pooled standard deviation was used for standardization.

Discussion

The primary aim of this study was to provide validity evidence for the MSCEIT by showing that it shared (a) meaningful structural relations with measures from emotions research, and (b) moderate relations with intelligence constructs (Gf and Gc). Generally, structural findings were consistent with previous reports and with the theoretical model of EI: Strategic EI (and particularly its Understanding component) was related to other measures of intelligence, meeting an important criterion for demonstrating this as a form of intelligence. Quite troubling, given the apparent similarity in the construct definitions for perceiving emotions in both areas, is that the Emotion Perception component of the MSCEIT and tests from emotions research are factorially distinct. Somewhat surprisingly, the emotions research measures loaded on an intelligence factor.

Empirical support for the two-area, four-branch theoretical model of EI was mixed. Exploratory and confirmatory factor analyses provided some support for the distinction between Experiential and Strategic EI, but not for the four-branch distinction at a lower level of a higher-order EI model. This is consistent with prior studies finding very high correlations between factors defined by the tests of the two Experiential EI branches (Day & Carroll, 2004). Overall, this outcome points to the possibility that the branches as measured with the MSCEIT are empirically hard to distinguish.

Perhaps the most surprising finding of the present study was that tests from emotions research (i.e., JACBART, Vocal-I) could not be subsumed under the MSCEIT tests of Emotion Perception. Despite the fact that the measurement intentions associated with the tests from the two research areas are highly similar, if not identical, they did not perform as indicators of the same common underlying factors. Given that there is a considerable research basis, both conceptual and empirical, behind the emotions research measures, this finding could be regarded as calling into question the validity of at least the Emotion Perception factor assessed by the MSCEIT.

In defense of the MSCEIT, different measures of interpersonal sensitivity have been found to exhibit near zero to moderate positive correlations (see Hall & Bernieri, 2001). Conceivably, such a finding might undermine the expectation of a substantial correlation between the MSCEIT's Emotion Perception and the emotions research measures. However, for the validation of the Emotion Perception branch measured with the MSCEIT, the focus should be placed on emotion recognition instead of the more general class of nonverbal sensitivity measures (e.g., the Profile of Nonverbal Sensitivity, PONS; Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979). In fact, there is evidence available for common variance between emotion recognition measures and this is true even for measures using different modalities, like verbal and facial expressions of emotions ($r = .24$, $N = 1264$, $p < .01$, see Scherer,

Table 2
Standardized Factor Loadings and Factor Correlations from Exploratory and Confirmatory Factor Analysis of MSCEIT, Gc, Gf, and Emotions Research Measures

	Exploratory factor analysis				Confirmatory factor analysis		
Indicators	F1	F2	F3	h^2	F1	F2	F3
Faces	.61**	.05	−.08	.37	.56		
Pictures	.60**	−.25**	.09	.38	.70	−.34	
Facilitation	.60**	.11	−.01	.39	.59		
Sensations	.30**	.26**	.07	.21	.34	.23	
Changes	.28**	.36**	.27**	.40	.27	.33	
Blends	.16	.40**	.35**	.44		.45	.34
Management	.06	.50**	.07	.29		.55	.26
Relations	−.03	.90**	−.01	.80		.85	
Vocabulary	−.10	−.17	.39**	.14			.24
Analogies	.06	−.02	.65**	.43			.68
Swaps	−.02	.06	.38**	.15			.42
Matrices	−.03	.09	.21	.06	—	—	—
Vocal-I	−.16	.05	.47**	.25			.43
JACBART	.02	.08	.35**	.15			.40
Factor correlations							
F2	.19*				.34		
F3	.19*	.31**			.28	.33	

Note. MSCEIT = Mayer-Salovey-Caruso Emotional Intelligence Test; Gc = crystallized intelligence; Gf = fluid intelligence; F = factor; h^2 = communalities; JACBART = Japanese and Caucasian Brief Affect Recognition Test; all loadings of the CFA (confirmatory factor analysis) model are significant at the .05-level. The same is true for the factor correlations in the CFA. Fit statistics for the CFA model are as follows: $\chi^2 = 49.35$, $df = 57$, $p = .754$; RMSEA = .00, CFI = 1.00.
* $p < .05$. ** $p < .01$.

in press). Although the measures used in the present study (Vocal-I and JACBART) showed only a small to medium intercorrelation ($r = .17$, $p < .05$; see Appendix), it can be stated that the pattern of correlations between these two measures and the MSCEIT subtests, as well as the factor-analytic results, fail to support the validity of the Emotion Perception measures of the MSCEIT.

The reasons for the lack of support for the validity of the MSCEIT found in this study are likely manifold. First, although there is overlap in the type of measurement procedures used to assess the ability to perceive emotions in faces in the research domains of EI and emotions research they also show apparent differences. The Faces subtest of the MSCEIT, which most closely resembles the type of measures used in emotions research (e.g., JACBART), has a smaller number of stimuli, it lacks the specificity of facial expressions of emotions typically found in measures in emotions research, and the instruction to rate the presence of a series of emotions is also different from the instruction used in the domain of emotions research. Measures in emotions research typically ask participants to select the primary emotion in a stimulus, rather than rate emotional intensity as in EI research. In addition, different types of stimuli are also used for some tasks (e.g., Pictures in the MSCEIT and auditory stimuli in the Vocal-I) that might represent test specifics not common to other indicators.

Second, the measures that originate from emotions research generally have predetermined correct or incorrect responses, while the MSCEIT subtests are scored consensually. It remains an open empirical question whether these different scoring methods converge or contribute to divergence in results for measures purportedly assessing the same latent variable.

Third, some of the measures emanating from emotions research contain a considerable component of speediness in administration procedures (e.g., JACBART). Again, it is not entirely clear how this might affect the results of the assessment procedures and correlations with external measures. Research on mental speed suggests, however, that this aspect of measurement might impact on the assessment of the target construct (see Danthiir, Roberts, Schulze, & Wilhelm, 2005). Because Gf and speed are substantially correlated (see Danthiir et al., 2005) this might also explain why the JACBART loaded on the same factor as intelligence markers (e.g., Swaps). The implication of this line of argument is that the emotions research measures, used a criterion for the MSCEIT in the present study, might themselves suffer from imperfect validity.

Overall, there appear to be a several methodological differences that are prevalent in different measurement approaches, which might contribute to the divergence of interindividual differences in ability estimates. Among these are differences in type of stimuli used, instructions and corresponding response options, scoring procedures, and the speeded component underlying these tests. Hence, although it can be hypothesized that at the conceptual level there is a single construct representing the ability to perceive emotions, which is a cause for individual differences in the measures used in EI and emotions research, specifics at the measurement level may cause a test of a corresponding measurement model to fail.

In light of this fact and the results reported in the present paper, it seems reasonable to expand the current research. For example, in future research of this kind, one might use a still broader battery of emotion perception measures (e.g., Ekman's standardized database of facial expressions, Ekman, 1973; the Diagnostic Assessment of Non Verbal Affect–Adult Facial Expressions, DANVA2-AF, Nowicki & Carton, 1993; the DANVA2-AP, Baum & Nowicki,

1998). Such a design would allow for a test of a measurement model with a single latent variable, namely Emotion Perception, which explains the common variance of the measures in the battery. This would directly address the question of whether it makes sense to assume that all of the different emotion perception measures actually assess the same underlying ability. It is also worth noting that it would go beyond the current practice of trying to answer this question by inspecting the bivariate correlations between the measures under investigation.

The strategy to broaden the spectrum of indicators also appears sensible for the other branches of the EI model supposedly underlying the MSCEIT (see Mayer, Panter, Salovey, Caruso, & Sitarenios, 2005). This might include a range of experimental measures emanating from emotions research that are tied to understanding and management (O'Sullivan, in press), as well as paradigms making use of emergent methodologies, including the situational judgment, implicit association, and conditional reasoning paradigms (see, e.g., Roberts et al., 2005). Indeed, in the related field of social intelligence, recent research using such an approach has proven fruitful by making it possible to triangulate the core latent variables of this domain (see Weis & Süß, 2005). The conceptual similarities between social and emotional intelligence (Austin & Saklofske, 2005) support the expectation that future research in this direction might yield valuable insights into the dimensional structure of EI.

To summarize, we are left with serious doubts concerning the construct validity of at least the assessment of Emotion Perception with the MSCEIT. Since validation is a cumulative enterprise, we cannot completely discount this EI measure at this stage. Further research will have to replicate the results presented here and extend the approach to other subscales of the MSCEIT. Nevertheless, at this stage we can not suggest that the MSCEIT has stood a test of an important aspect of the validation process.

References

Austin, E. J., & Saklofske, D. H. (2005). Far too many intelligences? On the communalities and differences between social, practical, and emotional intelligence. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 107—128). Cambridge, MA: Hogrefe & Huber.

Baum, K. M., & Nowicki, S., Jr. (1998). Perception of emotion: Measuring decoding accuracy of adult prosodic cues varying in intensity. *Journal of Nonverbal Behavior, 22,* 89–107.

Brackett, M. A., & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin, 29,* 1147–1158.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36,* 111–150.

Browne, M. W., Cudeck, R., Tateneni, K., & Mels, G. (2004). CEFA: Comprehensive Exploratory Factor Analysis, Version 2.00 [Computer software and manual]. Retrieved July 12, 2005, from http://quantrm2.psy.ohio-state.edu/browne/

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Coffey, E., Berenbaum, H., & Kerns, J. G. (2003). The dimensions of emotional intelligence, alexithymia, and mood awareness: Associations with personality, and performance on an emotional Stroop task. *Cognition and Emotion, 17,* 671–679.

Danthiir, V., Roberts, R. D., Schulze, R., & Wilhelm, O. (2005). Mental speed: On frameworks, fissures, and a platform for the future. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 27–46). Thousand Oaks, CA: Sage.

Davis, M. H., & Kraus, L. A. (1997). Personality and empathic accuracy. In W. Ickes (Ed.), *Empathic accuracy* (pp. 144–168). New York: Guilford.

Day, A. L., & Carroll, S. A. (2004). Using an ability-based measure of emotional intelligence to predict individual performance, group performance, and group citizenship behaviors. *Personality and Individual Differences, 36,* 1443–1458.

Ekman, P. (1973). *Darwin and facial expressions.* New York: Academic Press.

Hall, J. A., & Bernieri, F. J. (Eds.). (2001). *Interpersonal sensitivity: Theory and measurement.* Mahwah, NJ: Erlbaum.

Kafetsios, K. (2004). Attachment and emotional intelligence abilities across the life course. *Personality and Individual Differences, 37,* 129–145.

Legree, P. J., Psotka, J., Tremble, T., & Bourne, D. R. (2005). Using consensus based measurement to assess emotional intelligence. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 155–179). Cambridge, MA: Hogrefe & Huber.

Lopes, P. N., Salovey, P., & Straus, R. (2003). Emotional intelligence, personality, and the perceived quality of social relationships. *Personality and Individual Differences, 35,* 641–658.

MacCann, C. (2005). [New approaches to the assessment of emotional intelligence]. Unpublished raw data.

MacCann, C., Roberts, R. D., Matthews, G., & Zeidner, M. (2004). Consensus scoring and empirical option weighting of performance-based emotional intelligence (EI) tests. *Personality and Individual Differences, 36,* 645–662.

Markham, P. (2002). *Construction and validation of a tacit knowledge inventory for university students.* Unpublished honours dissertation, University of Sydney, Sydney, Australia.

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., et al. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior, 24,* 179–209.

Matthews, G., Roberts, R. D., & Zeidner, M. (2004). Seven myths about emotional intelligence. *Psychological Inquiry, 15,* 179–196.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science and myth.* Boston: MIT Press.

Mayer, J. D., Panter, A. T., Salovey, P., Caruso, D. R., & Sitarenios, G. (2005). A discrepancy in analyses of the MSCEIT—Resolving the mystery and understanding its implications: A reply to Gignac (2005). *Emotion, 5,* 236–237.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–34). New York: Basic Books.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2002). *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT): User's manual.* Toronto, Canada: Multi-Health Systems.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion, 3,* 97–105.

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the emotional Stroop task. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 216,* 1595–1607.

Nowicki, S., Jr., & Carton, J. (1993). The measurement of emotional intensity from facial expressions. *Journal of Social Psychology, 133,* 749–750.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, and Computers, 32,* 396–402.

O'Sullivan, M. (1982). Measuring the ability to recognize facial expres-

sions of emotion. In P. Ekman (Ed.), *Emotion in the human face* (pp. 281–314). New York: Cambridge University Press.

O'Sullivan, M. (in press). Trolling for trout, trawling for tuna: The methodological morass in measuring emotional intelligence. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *Emotional intelligence: Knowns and unknowns.* Oxford, UK: Oxford University Press.

Roberts, R. D., Schulze, R., Zeidner, M., & Matthews, G. (2005). Understanding, measuring, and applying emotional intelligence: What have we learned? What have we missed? In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 311–341). Cambridge, MA: Hogrefe & Huber.

Roberts, R. D., & Stankov, L. (1999). Individual differences in speed of information processing and human cognitive abilities: Toward a taxonomic model. *Learning and Individual Differences, 11,* 1–120.

Roberts, R. D., Zeidner, M., & Matthews, G. (2001). Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusions. *Emotion, 1,* 196–231.

Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. (1979). *Sensitivity to nonverbal communications: The PONS test.* Baltimore: Johns Hopkins University Press.

Scherer, K. (in press). Cognitive models of emotion: Synthesis and implications for emotional competence. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *Emotional intelligence: Knowns and unknowns.* Oxford, UK: Oxford University Press.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion interferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32,* 76–92.

Weis, S., & Süß, H-M. (2005). Social intelligence—A review and critical discussion of measurement concepts. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 203–230). Cambridge, MA: Hogrefe & Huber.

Wilhelm, O. (2005). Measures of emotional intelligence: Practice and standards. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 131–154). Cambridge, MA: Hogrefe & Huber.

Zeidner, M. Shani-Zinovich, I., Matthews, G., & Roberts, R. D. (2005). Assessing emotional intelligence in gifted and non-gifted high school students: Outcomes depend on the measure. *Intelligence, 33,* 369–391.

Appendix

Correlations Between MSCEIT Tasks, Gc and Gf Measures, and Measures From Emotion Research

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Faces	.07													
2. Pictures	.34	.08												
3. Facilitation	.36	.33	.08											
4. Sensations	.24	.22	.17	.06										
5. Changes	.23	.12	.32	.31	.08									
6. Blends	.14	.11	.26	.19	.43	.08								
7. Management	.14	−.06	.22	.18	.23	.34	.05							
8. Relations	.11	−.11	.16	.31	.43	.48	.47	.08						
9. Vocabulary	−.07	.04	−.05	−.05	.02	.11	−.02	−.05	.14					
10. Analogies	.08	.10	.12	.14	.30	.35	.13	.16	.20	.16				
11. Swaps	.04	.06	−.01	.15	.14	.11	.13	.16	.01	.33	.22			
12. Matrices	−.01	.00	.01	.13	.09	.20	.09	.14	.19	.16	.02	.18		
13. Vocal-I	−.10	−.01	−.06	.08	.22	.24	.16	.14	.20	.25	.20	.00	.10	
14. JACBART	.08	−.02	.12	.10	.20	.16	.19	.16	.09	.24	.27	.13	.17	.11

Note. MSCEIT = Mayer-Salovey-Caruso Emotional Intelligence Test; Gc = crystallized intelligence; Gf = fluid intelligence; JACBART = Japanese and Caucasian Brief Affect Recognition Test; *SD*s are given on the main diagonal of the matrix.

Received August 1, 2005
Revision received March 22, 2006
Accepted April 20, 2006 ■